COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 1055

FORWARD EXCHANGE MARKET UNBIASEDNESS: THE CASE OF THE AUSTRALIAN DOLLAR SINCE 1984

Peter C. B. Phillips and James W. McFarland

August 1993

FORWARD EXCHANGE MARKET UNBIASEDNESS: THE CASE OF THE AUSTRALIAN DOLLAR SINCE 1984[1]

Peter C. B. Phillips
Cowles Foundation for Research in Economics
Yale University

and

James W. McFarland
A. B. Freeman School of Business
Tulane University

Submitted: August, 1995
Revised: June, 1996

[1]This paper is dedicated to the memory of our dear friend and colleague Patrick McMahon, with whom we first began our work on this topic. Computations and graphics reported in the paper were performed by PCBP in programs written in GAUSS. Many of these programs are available in the software package COINT 2.0 of Ouliaris and Phillips (1993). The first author thanks the NSF for research support under SES 91-22142 and SBR 94–22922 and Tulane University for its hospitality. The authors both thank Sam Ouliaris for supplying the data, and Glena Ames for assistance in keyboarding the manuscript.

Abstract

This paper implements a robust statistical approach to regression with nonstationary time series. The methods were recently developed in other work, and are briefly exposited here. They allow us to perform regressions in levels with nonstationary time series data, they accommodate data distributions with heavy tails and they permit serial dependence and temporal heterogeneity of unknown form in the equation errors. With these features the methods are well suited to applications with frequently sampled exchange rate data, which generally display all of these empirical characteristics. Our application here is to daily data on spot and forward exchange rates between the Australian and US dollars over the period 1984–1991 following the deregulation of the Australian foreign exchange market. We find big differences between the robust and the non-robust regression outcomes and in the associated statistical tests of the hypothesis that the forward rate is an unbiased predictor of the future spot rate. The robust tests reject the unbiasedness hypothesis but still give the forward rate an important role as a predictor of the future spot rate.

1 INTRODUCTION

This paper applies some new econometric methodology to test the hypothesis that forward exchange rates are optimal predictors of future spot rates. The forward market unbiasedness hypothesis has been the subject of recent research that makes use of nonstationary time series methods. Some of this work has involved direct nonstationary regression techniques to fit empirical relationships between forward rates and future spot rates. In particular, Corbae, Lim and Ouliaris (1993), Moore (1992), and Phillips *et al.* (1996) have employed the Phillips–Hansen (1990) fully modified estimator and Park's (1992) canonical cointegrating regression technique in such regressions. Others have used error correction and reduced rank regression approaches (see Hakkio and Rush, 1989; and Baillie and Bolleslev, 1989). While these techniques cope with nonstationary data, they were not designed to deal with data that display strong outlier activity as well as nonstationarity, like exchange rates. There is good reason therefore to question whether the empirical results from the use of such techniques are statistically robust. The methods of this paper provide an avenue for addressing this question and thereby sorting out which effects may be due to nonstationarity and which to outlier behavior.

Let s_t be the natural logarithm of the spot exchange rate (S_t) of a given currency in terms of the US dollar at time t, and let $f_{t,k}$ be the natural logarithm of the forward exchange rate ($F_{t,k}$) of the currency contracted at time t for delivery at time $t+k$. Forward market unbiasedness tests can be based directly on the regression model

$$s_{t+k} = \alpha + \beta f_{t,k} + u_{t+k} \tag{1}$$

and the hypothesis is couched in the parametric form

$$H_0 \ : \ \alpha = 0 \ , \quad \beta = 1 \ . \tag{2}$$

Much conventional empirical methodology in this field is motivated by an attempt to avoid working with the nonstationary levels data s_{t+k} and $f_{t,k}$. To achieve this, (1) and (2) are often combined under H_0 to produce a model whose dependent variable is the spot return $s_{t+k} - s_t$ and whose regressor is the forward premium $f_{t,k} - s_t$, as in the following equation

$$s_{t+k} - s_t = \alpha + \beta \left(f_{t,k} - s_t \right) + u_{t+k}.$$

Return/premium regressions such as this are studied in Phillips and McFarland (1996). It is shown there that such regressions suffer some major disadvantages in comparison with (1). For instance, they lead to a lower rate of estimator convergence under the null hypothesis; and, when $\beta \neq 1$, the OLS estimator of β converges in probability to zero, even when the forward rate has predictive content in the original model (*ie* when $\beta \neq 0$ in (1)). Thus, the potential for spurious inference is high in such regressions and it is preferable to work with the levels regression formulation (1) instead. In some earlier joint work (see Goodhart *et al.,* 1992), Patrick McMahon pointed to weaknesses in return/premium regressions and used subsample estimation to reveal the unreliability of the coefficients in these regressions.

When the contract period k exceeds the sampling interval (the case of overlapping data) — as will be the case for daily data — the error u_{t+k} in (1) can be expected to be serially dependent and possibly temporally heterogeneous. Moreover, since s_t and f_{t+k} typically display stochastic nonstationarity in their log-levels form, (1) involves nonstationary data as well as serially dependent errors. The fully modified least squares (FM–OLS) procedure of Phillips–Hansen (1990) accommodates both these characteristics of the regression model (1) and this is why it has been used in recent research. However, like all least squares regression techniques, FM–OLS was not designed to deal specifically with data (like exchange rates) where there is prominent outlier behavior.

The methodology we use in this paper to estimate regression equations such as (1) is designed to be statistically robust to outlier activity in the regressors and the errors. Our procedure is called fully modified least absolute deviations (FM–LAD) and has the same robust regression features as the least absolute deviation (LAD) regression estimator (see Basset and Koenker, 1978). Like FM–OLS, the FM–LAD estimator also corrects for endogeneity and serial correlation and is therefore well suited to forward market exchange models such as (1) which embody all of these characteristics. The asymptotic theory of the FM–LAD estimator is derived in Phillips (1995) and the estimator has been successfully applied to some historical data series that relate to the 1920's experiment with floating exchange rates (see Phillips, McFarland and McMahon, 1996). We will describe the main features of the method here in Section 2.

The present application is to daily exchange rate data for the Australian dollar in the period following the financial deregulation of the early 1980's. The data and its time series and tail slope characteristics are analyzed in Section 3; model selection exercises including some recursive cointegration analyses are conducted in Section 4; our empirical regression results and tests of unbiasedness are reported in Section 5; nonparametric estimates of the probability densities of the equation errors and exchange rate returns are computed in Section 6; and some brief concluding comments are made in Section 7.

2 FM–LAD ESTIMATION AND TESTING

The FM–LAD estimation procedure was developed in Phillips (1995). Here, we briefly describe the construction of the FM–LAD estimator, its associated test statistics and the relevant asymptotic theory. It will suffice for this exposition and our intended application to use the simple cointegrated system

$$y_t = \beta' x_t + u_{0t} \tag{3a}$$

$$\Delta x_t = u_{xt} \tag{3b}$$

where $u_t' = (u_{0t},\, u_{xt}')$ is a stationary m-vector time series ($m = 1 + m_x$) with spectral density matrix $f_{uu}(\lambda)$. The long-run covariance matrix of u_t is

$$\Omega_{uu} = 2\pi f_{uu}(0) = \left[\begin{array}{cc} \Omega_{00} & \Omega_{0x} \\ \Omega_{x0} & \Omega_{xx} \end{array} \right], \tag{3}$$

where the partition is conformable with that of the vector u_t. We assume that $\Omega_{xx} > 0$ (*i.e.*, Ω_{xx} is positive definite), so that x_t in (3a) is a full rank I(1) process in the sense that the number of unit roots in the stochastic process x_t is equal to m_x, the dimension of x_t. In cases where u_t does not possess finite second moments the matrix Ω_{uu} in (4) is not well-defined. However, it is still possible in such cases to construct a pseudo long-run variance of u_t, as discussed in Phillips (1990, p. 51), and we will proceed as if this has been done.

In FM–LAD estimation we also need to work with the transformed error $v_t = sign(u_{0t})$, and we therefore define the long-run covariance matrix of $w_t' = (v_t, u_{xt}')$ as

$$\Omega_{ww} = 2\pi f_{ww}(0) = \begin{bmatrix} \Omega_{vv} & \Omega_{vx} \\ \Omega_{xv} & \Omega_{xx} \end{bmatrix} ,$$

which is partitioned conformably with w_t. Note that v_t is bounded and has finite moments of all orders. But this is not true of u_{xt} and in cases where the second moments of u_{xt} do not exist we may again employ a pseudo-variance interpretation of Ω_{ww}.

The LAD estimator of β in (3a) is the extremum estimator

$$\beta_{\text{LAD}} = \arg\min_{\beta}\{n^{-1}\Sigma_1^n |y_t - x_t'\beta|\} . \tag{4}$$

When the regressors x_t are fixed this estimator has an asymptotic normal distribution and is $\sqrt{n}$ consistent for β in (3a). When x_t is an I(1) process and (3b) holds, this asymptotic theory no longer applies in general. Instead, the LAD estimator, just like OLS, suffers from bias and non-scale nuisance parameter problems even in the limit as $n \to \infty$.

The FM–LAD estimator is designed to address the difficulties that are encountered by the LAD estimator while at the same time retaining its robustness features with regard to heavy tailed errors. As with the FM–OLS estimator, we modify LAD to account for possible endogeneities in the x_t regressor variables and serial dependence in the errors. The FM–LAD estimator is defined by

$$\beta_{\text{LAD}}^{+} = \beta_{\text{LAD}} - \left(\tfrac{1}{2}\hat{f}(0)\right)(X'X)^{-1}(X'\Delta X - n\widehat{\Delta}_{xv}^{+}) , \tag{5}$$

where $X'X = \Sigma_1^n x_t x_t'$, $X'\Delta X = \Sigma_1^n x_t \Delta x_t'$, and $\hat{f}(0)$ is a consistent estimator of the probability density of u_{0t} at the origin.

The matrix $\widehat{\Delta}_{xv}^{+}$ in (6) is a consistent estimator of the one-sided long-run covariance matrix

$$\Delta_{xv}^{+} = \Sigma_{k=0}^{\infty} E(u_{x0} v_k^{+}) , \tag{6}$$

where

$$v_t^{+} = v_t - \Omega_{vx}\Omega_{xx}^{-1}\Delta x_t , \tag{7}$$

and

$$v_t = \text{sign}(u_{0t}) . \tag{8}$$

To estimate Δ_{xv}^{+} we need first to estimate the modified error v_t^{+}, and this involves the estimation of v_t. To do so, we run a first stage LAD regression which produces the error estimate $\hat{u}_{0t} = y_t - \beta'_{\text{LAD}} x_t$. Setting $\hat{v}_t = \text{sign}(\hat{u}_{0t})$, we construct

$$\hat{v}_t^{+} = \hat{v}_t - \widehat{\Omega}_{\hat{v}x}\widehat{\Omega}_{xx}^{-1}\Delta x_t \ , \tag{9}$$

using conventional kernel estimates of the long-run covariance matrices Ω_{vx} and Ω_{xx}, whereupon we can estimate Δ_{xv}^{+} by using a kernel estimate of the one-sided long-run covariance of u_{xt} and $\widehat{v}_t^{+}$ (see Park and Phillips, 1988) . Note that we can write

$$\Delta_{xv}^{+} = \Delta_{xv} - \Delta_{xx}\Omega_{xx}^{-1}\Omega_{xv}, \text{ where } \Delta_{xv} = \Sigma_{k=0}^{\infty} E(u_{x0}v_k), \ \Delta_{xx} = \Sigma_{k=0}^{\infty} E(u_{x0}u'_{xk}), \tag{10}$$

so that the estimation of $\omega_{vv\cdot x} = \Omega_{vv} - \Omega_{vx}\Omega_{xx}^{-1}\Omega_{xv}$ effectively involves the estimation of the four submatrices Δ_{xv}, Δ_{xx}, Ω_{xx}, and Ω_{xv}. Once again, when variances are infinite we can employ pseudo-variance interpretations of these quantities and these pseudo-variances may all be estimated in the usual way with finite samples of data.

The asymptotic theory of the FM–LAD estimator β_{LAD}^{+} is given by

$$\beta_{\text{LAD}}^{+} - \beta \sim N\left(0, \ (1/2f(0))^2 \omega_{vv\cdot x}(X'X)^{-1}\right) \tag{11}$$

where $\widehat{\omega}_{vv\cdot x} = \widehat{\Omega}_{vv} - \widehat{\Omega}_{vx}\widehat{\Omega}_{xx}^{-1}\widehat{\Omega}_{xv}$. Hence, Wald statistics can be constructed in the usual way from this mixture normal approximation to test restrictions on the parameter vector β and such statistics have limiting chi-squared distributions with degrees of freedom equal to the number of restrictions.

When the system (3) has infinite variance errors, it is shown in Phillips (1995) that (11) still holds but with $\omega_{vv\cdot x} = \Omega_{vv}$. Thus, the FM–LAD estimator has an asymptotic mixed normal approximation whether or not the error variances in the underlying system are finite. So the FM–LAD estimator has the attractive feature that its limit theory in the case of cointegrated systems like (3) may be used for statistical inference irrespective of the tail thickness of the errors. This feature makes the estimator β_{LAD}^{+} and its associated Wald tests very useful in the context of non-stationary data with heavy tails.

3 THE DATA AND OUTLIER CHARACTERISTICS

The data employed in this study consist of daily spot, one-month forward and three-month forward exchange rates for the Australian dollar against the US dollar over the period beginning 3 January 1984 and ending 2 April 1991. There are 1,830 observations in total. This period follows the decision made in December 1983 by the Australian government to float the Australian dollar and to abolish a major part of existing exchange controls. Not all exchange controls were abolished at this time however and, in particular, trading banks alone retained the privilege of dealing in spot and forward exchange markets until 1988 when access to the foreign exchange market was given to nonbank licensed foreign exchange dealers by the Reserve Bank

of Australia. Juttner (1990, ch. 22) provides a recent history of institutional arrangements and exchange rate regimes in Australia and the reader is referred to this source for further details.

Figures 1 and 2 graph the levels of the series and the figures show the spot rate (SR) at $t+k$ (i.e., S_{t+k}) against the two forward rates (i.e., $F_{t,k}$) for the one-month and three-month contracts. The data are aligned so that the spot rate is shown for the date when the forward contract matures. As is apparent from both figures, the data behaves as if they have no fixed mean.

- - - - - - - - - - - - - - - - - - - -
Figures 1 and 2 about here
- - - - - - - - - - - - - - - - - - - -

Figures 3 and 4 show the spot rate and forward rate return series (i.e., $\Delta s_t = \ln S_t - \ln S_{t-1}$ and $\Delta f_{t,k} = \ln F_t - \ln F_{t-1}$) over the same historical period. Both these graphs show evidence of some large outlier activity, especially the one-month forward rate over 1989–1991.

- - - - - - - - - - - - - - - - - - - -
Figures 3 and 4 about here
- - - - - - - - - - - - - - - - - - - -

The outlier activity in the data can be studied by estimating the tail slope parameter (α, say) of the data distribution, using order statistic methods as explained in Phillips, MacFarland and McMahon (1996). These methods rely on the use of the s largest order statistics of the exchange returns in each tail. The choice of s can be data-based (as described in Hall and Welsh, 1985), but this relies on the further assumption that the tails are asymptotically of the Pareto form and can be expanded in powers of $x^{-\alpha}$. While this further assumption is restrictive, it does enable us to provide entirely data-determined (or adaptive) estimates of the tail slope. Estimates and adaptive estimates of α for our data are given below in Table 1.

Table 1
Point Estimates of Tail Slope Parameter for Exchange Rate Returns

Series	s	Left tail			Right tail		
	40	2.516		(0.397)	3.769		(0.595)
	50	2.450		(0.346)	3.405		(0.481)
Spot rate	75	2.214		(0.255)	2.829		(0.326)
	100	2.011		(0.201)	2.741		(0.274)
	$\widehat{s}$	2.452	[58]	(0.322)	3.377	[58]	(0.443)
	40	2.457		(0.388)	2.878		(0.455)
	50	2.499		(0.353)	3.005		(0.424)
1-month forward rate	75	2.680		(0.309)	2.658		(0.306)
	100	2.471		(0.247)	2.674		(0.267)
	$\widehat{s}$	2.661	[73]	(0.311)	2.568	[83]	(0.281)
	40	2.877		(0.454)	4.076		(0.644)
	50	3.033		(0.429)	3.925		(0.555)
3-month forward rate	75	2.747		(0.317)	3.076		(0.355)
	100	2.769		(0.276)	3.082		(0.308)
	$\widehat{s}$	2.856	[70]	(0.341)	3.023	[76]	(0.346)

[ ] = $\widehat{s}$, adaptive estimate of order statistic number.
( ) = standard error of $\hat{\alpha}_{s.}$

There is evidence from the estimates given in Table 1 that spot exchange rate returns are asymmetrically distributed with heavier left tails than right tails. The left tail slope coefficient estimate $\hat{\alpha}_{\hat{s}} = 2.452$ is 2.87 standard deviations from the right tail slope coefficient estimate $\hat{\alpha}_{\hat{s}} = 3.377$. The asymmetry in tail slope is also apparent from the graphed return series shown in Figure 3, where we see large negative outliers in 1985, 1986, 1989 and 1990 and smaller positive outliers for the same years.

The one-month forward rate return series has right and left tail slope estimates that are much more balanced. There is no evidence of asymmetry in this series. But the slope estimates show strong evidence of thick tailed distributions. The adaptive estimates $\hat{\alpha}_{\hat{s}}$ are more than one asymptotic standard deviation less than 3.0 and more than two standard deviations less than 3.30. The three-month forward rate tail slope estimates are a little higher ($\hat{\alpha}_{\hat{s}} = 2.856$ for the left tail and $\hat{\alpha}_{\hat{s}} = 3.023$ for the right tail). These are still significantly lower than 4.0, so that there is doubt whether a finite fourth moment distribution is an appropriate model for the data and strong evidence in each case against Gaussianity.

4 MODEL SELECTION, UNIT ROOTS AND COINTEGRATION

We first set out to find the most suitable model for each of the series in the autoregressive-moving average (ARMA) with deterministic trend (Tr) class. This process not only helps to characterize the time series features of the data but also enables us to address

the question of whether the data is better modelled with a deterministic trend and whether there is a unit root in the model. Our methodology in this exercise follows that of Phillips and Ploberger (1994). Specifically, we set up a general reference model in the ARMA(p,q) + Tr(t) class of the form

$$y_s = \alpha y_{s-1} + \Sigma_{i=1}^{p-1}\varphi_i \Delta y_{s-i} + \Sigma_{j=1}^{q}\psi_j \varepsilon_{s-j} + \Sigma_{j=0}^{t} b_j s^j + \varepsilon_s \ , \tag{12}$$

for each series and used the model selection algorithm of Phillips and Ploberger (1993) to empirically determine the trend degree ($\widehat{t}$) and ARMA orders ($\hat{p},\hat{q}$). The version of that algorithm that we employ here makes use of: (i) the BIC criterion (Schwarz, 1978) to select the lag orders and the trend degree; and (ii) recursive least squares estimation of (19) along the lines of Hannan–Rissanen (1982). In addition, we evaluated the model that was chosen and estimated this way with the same model estimated with an autoregressive unit root imposed.

Table 2
Model Selection

Variable (log-levels)	Trend degree	ARMA(p,q) orders		Long-run AR coefficient	PIC odds in favor of a unit root
	$\widehat{t}$	$\hat{p}$	$\hat{q}$	$\widehat{a}$	
Spot Rate	–1	1	1	0.999	1765
1-month Forward Rate	–1	1	1	0.999	1553
3-month Forward Rate	–1	1	0	0.999	1798

The results of this model selection exercise are shown in Table 2 and Figures 5-6. The chosen trend degree in each case is $\widehat{t} = -1$, so that a model of the form (19) is selected with no intercept or trend. The algorithm chooses an AR(1) model for the spot rate and the three-month forward rate but chooses an ARMA(1, 1) for the one-month forward rate. In the latter case the fitted model is (with t-ratios in parentheses):

$$y_s = \underset{(1596.05)}{0.999}\ y_{s-1} - \underset{(-3.736)}{0.087}\ \varepsilon_{s-1} \ .$$

The moving average effect is small, negative but quite significant in this case. The long run AR coefficient (the parameter a in model (19)) is, in fact, unity to the fourth decimal place, as it is for the other series (although the fourth digit is not shown in Table 2). When the algorithm is restricted to choose from autoregressive models the selected model is an AR(2) with the following fitted form

$$y_s = \underset{(1597.15)}{0.999}\ y_{s-1} - \underset{(-3.566)}{0.0832}\ y_{s-2} \ .$$

The final column in Table 2 gives the value of the posterior information criterion (PIC) in favor of a unit root. The PIC criterion is derived in Phillips and Ploberger (1994, 1995) and compares the ARMA$(\hat{p},\hat{q})$ + Tr$(\widehat{t})$ model for y_s with the ARMA$(\hat{p}-1,\hat{q})$ + Tr$(\widehat{t})$ model for Δy_s. In each case the criterion favors the model with a unit root present.

- - - - - - - - - - - - - - - - - - - -
Figures 5-6 about here
- - - - - - - - - - - - - - - - - - - -

Figures 5 and 6 give model selection graphics (for the spot rate and one-month forward rate) which help to show how well determined the selected model is in the given class. Fig. 5 displays the surface of PIC values against an array of (p, t) values for an AR(p) + Tr(t) model with $p \leq 5$ and $t \leq 1$. Since we choose the model with the highest PIC value (i.e., highest likelihood odds in its favor) it is clear from Fig. 5 that $(p, t) = (1, -1)$ is quite well determined for the spot rate. For the one-month forward rate, we see from Fig. 6 that the choice is $(p, t) = (2, -1)$, and again this is quite well determined.

We also performed model selection exercises in vector autoregressions for jointly modelling the spot rate and forward rate. The model is a vector version of (12) with no moving average component and has the form

$$\Delta y_s = A y_{s-1} + \Sigma_{i=1}^{p-1} \Phi_i \Delta y_{s-i} + \Sigma_{j=0}^{t} b_j s^j + \varepsilon_s \ , \quad s = 1, ..., T \tag{13}$$

with the levels coefficient matrix A taking the possibly reduced rank form

$$A = \alpha \beta' \ , \tag{14}$$

for certain $n \times r$ matrices α and β of rank r (here $n = 2$ and $0 \leq r \leq 2$).

When the rank $r = 1$ in (14), there is cointegration between the components of the integrated time series y_t. However, since the autoregressive lag length p in (13) is also unknown *a priori*, the problem of determining the rank r cannot be separated from the problem of order selection. Indeed, as argued recently in Phillips (1996), there is good reason to consider the cointegrating rank as an order selection problem in itself. As such, it is possible to treat (r, p) as a pair of parameters which can be jointly determined by order selection techniques. This approach was taken in the present paper, and we computed BIC values for an array of (r, p) values for the fitted model (13) using the formula

$$\text{BIC}(r, p) = \ln |\widehat{\Sigma}(r, p)| + \{n^2(p-1) + 2nr - r^2\} \ln(T)/T, \tag{15}$$

where $\widehat{\Sigma}(r, p)$ is the residual error covariance matrix from a fitted reduced rank regression of the form (13) with A matrix of rank r and p autoregressive lags.

Figures 7-8 provide graphs of the results of this joint model selection exercise for two VAR systems: one with the one-month forward rate and the other with the three-month forward rate. In each case Figure (a) shows the surface of BIC values for an (r, p) array with $0 \leq r \leq 2$ and $1 \leq p \leq 6$. (The BIC value for an equation with no autoregressive lag is so large that it distorts the scale of the surface and is therefore omitted — clearly $\hat{p} \geq 1$). Also Figure (b) shows a cross section of the surface to exhibit the BIC values for $0 \leq r \leq 2$, given the chosen autoregressive lag $\hat{p}$. For the two systems, the BIC criterion (15), with $(\hat{r}, \hat{p}) = \arg\min \text{ BIC}(r, p)$, leads to the same choice of $\hat{r} = 1$, $\hat{p} = 1$. The lag length seems in both cases to

be a little better determined than the cointegrating rank. Interestingly, this is true for the system involving the spot rate and the one-month forward rate (Figures 7(a) and (b)), where an AR(2) was earlier selected for the one-month forward rate as a univariate series (refer to Table 2).

- - - - - - - - - - - - - - - - - - -
Figures 7- 8 about here
- - - - - - - - - - - - - - - - - - -

Tests were conducted to assess the adequacy of the cointegrating regression model (1) over the sample period. The residual based Z_a and Z_t tests of Phillips and Ouliaris (1990) are reported in Table 3. The tests were computed using: (i) a fixed lag (set at $\ell = 10$) long-run variance estimator; (ii) a data-based long-run variance estimator based on the Andrews–Monahan (1992) AR prefiltered and recolored procedure; and (iii) the data-based long-run variance estimator of Lee and Phillips (1993) that uses ARMA model selected prefiltering and recoloring. Similarly, the ADF tests used a fixed AR lag length ($p = 10$) and a data-based (BIC) AR order selector $\hat{p}$. All of the tests confirm that the relationship (1) is cointegrating for the two forward contract periods, although the evidence is marginal in the case of the three-month relationship. The 5% critical values given in Table 3 are from the COINT 2.0 regression package (Ouliaris and Phillips, 1993). Table 3 also includes the results of the Phillips and Ploberger (1994, 1995) PIC unit root test applied to the residuals from the regression (1) for each currency. This test involves a data-based model selection procedure prior to the construction of the odds ratio. The odds in favor of cointegration are $4.8 \times 10^7 : 1$ in the case of the one-month forward rate relationship and 1.055:1 in the case of the three-month forward rate relationship. These outcomes corroborate the conclusions of the residual based tests, including the strong (marginal) conclusion in the one-month (three-month) relationship.

Table 3
Residual Based Tests of Cointegration

Equation	$\ell = 10$	Data-based $\hat{\ell}$	$\ell = 10$	Data-based $\hat{\ell}$	$p = 10$	Data-based $\hat{p}$	PIC odds in favor of cointegration
	Z_a	Z_a	Z_t	Z_t	ADF	ADF	
1-month Forward Contract	−103	−86.184	−7.300	−6.663	−4.876	−6.511	-4.88×10^7
3-month Forward Contract	−26.093	−22.690	−3.704	−4.280	−4.280	−3.512	1.055
5% critical values (cv's):	Z_a cv = −19.613;		Z_t cv = −3.465				

To exhibit the evidence in support of the cointegrating equation (1) over subperiods, the Z_a test was computed recursively over the historical period starting from the 100'th observation. The results for the one-month and three-month forward rate relationships are shown in Figures 9 and 10. The relationship between the spot rate and one-month forward rate is strongly supported by the nearly monotonic declining graph of the Z_a statistic in Figure 9. If equation (1) holds at this contract horizon,

this is exactly the behavior in the statistic that we would expect as the sample size grows and more information on the relationship accumulates, since Z_a diverges to $-\infty$ as the sample size $n \to \infty$ when (1) holds.

- - - - - - - - - - - - - - - - - -
Figures 9-10 about here
- - - - - - - - - - - - - - - - - -

We note a similar generally declining value in the recursive Z_a statistic for the three-month forward rate shown in Figure 10. But the evidence is not as strong and the recursive statistic is marked by subperiods, notably mid-1984 and 1987–1988, when the evidence shifted against the existence of a cointegrating relationship between the spot rate and three-month forward rate. As noted earlier, the period 1987–1988 involved some further important deregulation in the Australian foreign exchange market that admitted nonbank licensed participants to the market and the changes of this period may well have had a destabilizing influence on the relationship between the spot and forward rates over the longer contract horizon.

By the end of the sample period the evidence in favor of a cointegrating relationship in both cases seems quite strong. Since the time horizon is much longer in the case of the three-month forward contract we would naturally expect the forward rate in this case to be a less satisfactory predictor, as indeed is apparent from the data graphics in Figures 1 and 2. In addition, we might expect the residual in equation (1) to show longer temporal dependence in this case. To allow for this in the calculation of the recursive Z_a tests, we used a data-based long-run variance estimator, so that the lag truncation or bandwidth parameter naturally accommodated itself to the degree of temporal dependence in the data. Also the bandwidth parameter was selected on a period by period basis in an optimal data-based way as the recursive calculations proceeded, so that as the sample size grew the bandwidth parameter was adjusted accordingly.

5 ROBUST ESTIMATION AND TESTS OF FORWARD MARKET UNBIASEDNESS

Equation (1) was estimated for the one-month and three-month forward contract periods by OLS, reduced rank regression (RRR), FM–OLS, LAD and FM–LAD. The results are shown in Table 4. Standard errors, t-ratios and Wald statistics for testing the joint hypothesis $H_0 : \alpha = 0, \beta = 1$ are given in the table for the FM–OLS and FM–LAD procedures only. Note that conventional tests are not asymptotically valid in the case of the OLS and LAD estimators because of the nonstationarity and temporal dependence of the data for the reasons explained in Phillips–Durlauf (1986) and Park–Phillips (1988). However, the OLS and LAD coefficient estimates are both consistent, even though they suffer from second order bias, and it is therefore of some interest to report these estimates and determine the impact of the FM–OLS and FM–LAD estimator modifications.

The main empirical results are as follows:

(i) There are big differences between the FM–OLS and FM–LAD coefficient estimates, especially with regard to the slope coefficient β. For the equation involving the one-month forward rate we have $\beta^{+}_{\mathrm{LAD}} = 0.910$ and $\beta^{+}_{\mathrm{OLS}} = 0.980$, which differ by more than three FM–LAD asymptotic standard errors. For the three-month contract rate equation we have $\beta^{+}_{\mathrm{OLS}} = 0.883$, which is more than four FM–LAD asymptotic standard errors greater than the FM–LAD estimate $\beta^{+}_{\mathrm{LAD}} = 0.70$.

TABLE 4
Empirical Estimates of Equation (1)

Equation	Estimation Method	Parameters, standard errors and t-ratios						Joint Test
		α	s_α	$t_\alpha = \frac{\hat{\alpha}}{s_\alpha}$	β	s_β	$t_\beta = \frac{\hat{\beta}-1}{s_\beta}$	F
1-month Forward Contract	OLS	−0.023			0.904			
	RRR	0.001			0.998			
	FM–OLS	−0.002	0.009	−0.202	0.980	0.032	−0.593	1.518
	LAD	−0.022			0.894			
	FM–LAD	−0.017	0.006	−2.611 [b]	0.910	0.022	−4.035 [a]	29.307 [a]
3-month Forward Contract	OLS	−0.085			0.680			
	RRR	−0.004			0.934			
	FM–OLS	−0.025	0.029	−0.866	0.883	0.092	−1.256	2.485
	LAD	−0.077			0.681			
	FM–LAD	−0.071	0.012	−5.611 [c]	0.700	0.040	−7.381 [a]	70.838 [a]

Legend: (i) One tail: [a] significant at 0.1% level; (ii) Two tail [b] significant at 5% level [c] significant at 0.1% level

Note that the FM–OLS estimates are in both cases much closer to the value $\beta = 1$. This has a major impact on the inferred statistical properties of the risk premium

$$p_{t,k} = f_{t,k} - E_t(s_{t+k}) = (1-\beta)f_{t,k} - \alpha - E_t(u_{t+k}) \ .$$

for, if $\beta \neq 1$, this implies that the risk premium $p_{t,k}$ is nonstationary.

(ii) In testing the hypothesis $H_\beta : \ \beta = 1$ there are similar major differences between the procedures. For both contract periods H_β is clearly rejected by the data using the FM–LAD, but not rejected using FM–OLS. Thus, robust regression methods do seem to make an important difference in inference about this parameter. There are similar differences in inference with respect to α. In testing $H_\alpha : \ \alpha = 0$, the FM–LAD tests reject the hypothesis (strongly in the case of the three-month forward contract equation), whereas the hypothesis is not rejected using the FM–OLS test.

(iii) Tests of the joint hypothesis $H_0 : \ \alpha = 0$ and $\beta = 1$ also lead to big differences between the procedures. The FM-LAD chi-squared test statistic is large and significant beyond the 0.1% level for both the one-month contract and three-month contract equations. On the other hand, the FM–OLS test does not reject the hypothesis H_0. Again, the use of robust methods makes an important difference in the inferences that are drawn from the regression.

(iv) Looking at Table 4 we see that the biggest differences in the regression coefficient estimates arise between FM–OLS and the other three procedures. The endogeneity and serial correlation corrections lead to substantial changes in the OLS estimates, e.g., from $\widehat{\beta} = 0.680$ to $\beta^{+}_{\mathrm{OLS}} = 0.883$ in the case of the three-month contract equation. On the other hand, the LAD estimate is 0.681 and FM–LAD estimate is 0.700, both quite close to the OLS estimate. This reveals that the correction terms in the FM–OLS procedure carry much more weight and lead to much bigger changes in the estimate than they do for the LAD procedure. A possible explanation for this difference is that the LAD corrections rely on the residual function $\hat{v}_t = \mathrm{sign}(\hat{u}_{0t})$, which is bounded rather than the residual $\hat{u}_{0t}$ itself — see formula (6) and equations (7)–(11) above. The LAD corrections are therefore less likely to be affected by outliers in the equation errors than the OLS corrections that lead to the FM–OLS estimate.

(v) Table 4 also gives the RRR estimates of the coefficients for the two equations. These estimates are closest to the FM–OLS estimates and like FM–OLS are very different from the robust estimates. For the two equations considered, the RRR estimate of β is actually closer to unity than the FM–OLS estimate and this is particularly so in the three-month contract equation. These outcomes indicate that RRR estimation is very susceptible to the presence of outliers in the data. The results reported were obtained from a vector autoregression with six lags and very similar results were found when the lag length was varied from $p = 1$ to $p = 20$. So there does not appear to be much sensitivity to lag length in this case.

6 NONPARAMETRIC DENSITY ESTIMATES

Using the estimated residuals $\hat{u}_{0t}$ from the FM–OLS and FM–LAD regressions, together with the exchange rate returns $u_{xt} = \Delta x_t$ for one-month and three-month contracts, we computed kernel estimates of the probability density of u_{0t} and u_{xt}. A normal kernel was used in these computations in combination with a data-based ("plug in") optimal bandwidth based on Silverman's (1986, p. 45) recommendations.

Figures 11–12 graph these densities for the one-month forward contract equation. For comparison purposes, each of these figures also graphs a $N(0,\, s^2)$ density with variance s^2 equal to the sample variation of the data (either the residual $\hat{u}_{0t}$ or the return $u_{xt} = \Delta x_t$). The leptokurtosis and heavy tailed properties are evident in each of these nonparametric estimates and these features stand out clearly against the fitted normal curve. The heavy tail is especially marked in the case of the equation error density shown in Figure 11. Similar results were obtained for the three-month forward contract equation residuals and return data.

- - - - - - - - - - - - - - - - - -

Figures 11-14 about here

- - - - - - - - - - - - - - - - - -

In Figure 11 the equation errors u_{0t} are estimated using the FM–LAD residuals. In view of the important differences noted earlier between the FM–LAD and FM–

OLS coefficient estimates, we thought it useful to graph the nonparametric error densities based on the FM–OLS residuals as well. Figures 13 and 14 graph these error densities against those of the FM–LAD residuals for ease of comparison. The results are quite revealing. In both cases the FM–OLS residual density is noticeably skewed, with a mode that is substantially shifted away from the origin. The FM–LAD residual density is much less skewed and is better centered on the origin with a mode that is actually quite close to the origin. These estimated densities suggest that the corrections in the FM–OLS procedure are influenced by outliers in the data and result in a skewed residual density that compensates for the large correction that occurs in modifying OLS to FM–OLS. On the other hand, the FM–LAD estimate and fitted residuals do not seem to be strongly affected by outliers in the data.

This conclusion is corroborated by the data and regression line plots which are shown in Figures 15 and 16. The scatter plots of the data reveal some distinct outliers from a forward rate prediction model of the form (1). The fitted regression lines show how the FM–OLS coefficients are clearly influenced more by these outliers than FM–LAD. This is especially noticeable in the case of the three-month forward contract relation. These figures also show the regression lines fitted by reduced rank regression, which seem to be even more influenced by outliers in the data than the FM–OLS estimates.

- - - - - - - - - - - - - - - - - - -
Figures 15-16 about here
- - - - - - - - - - - - - - - - - - -

7 CONCLUSION

This paper reports an empirical application of robust nonstationary regression to the Australian foreign exchange market over the period 1984–1991. Both one-month and three-month forward exchange rates are found to have substantial predictive content for future spot rates. But, unlike optimal Gaussian regression techniques, the robust regression tests do not support the hypothesis that these forward rates are unbiased predictors of the future spot rate. Adaptive point estimates of the tail slope of the distributions of the equation errors and exchange returns, together with nonparametric estimates of the densities of these distributions all give strong evidence of heavy tails and support the use of robust, nonstationary regression methods.

References

[1] Andrews, D. W. K. and C. Monahan (1992). "An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator," *Econometrica*, 60, 953-966.

[2] Baillie, R. T. and T. Bollerslev (1989). "Common stochastic trends in a system of real exchange rates," *Journal of Finance*, 44, 167–181.

[3] Bassett, G. and R. Koenker (1978). "Asymptotic theory for least absolute error regression," *Journal of the American Statistical Association*, 73, 618–622.

[4] Baxter, M. (1994). "Real exchange rates and real interest rate differentials: Have we missed the business cycle relationship," *Journal of Monetary Economics, 33, 5-37.*

[5] Corbae, D., K–G Lim and S. Ouliaris (1993). "On cointegration and tests of forward market unbiasedness," *Review of Economics and Statistics*, 74, 728-732.

[6] Engle, R. F. and C. W. J. Granger (1987). "Cointegration and error-correction: Representation, estimation and testing," *Econometrica*, 55, 251–276.

[7] Goodhart, C. A. E., P. C. McMahon and Y. L. Ngama (1992). "Does the Forward Premium Discount help to predict the future change in the Exchange Rate?" mimeographed (University of Birmingham).

[8] Hakkio, C. S. and M. Rush (1989). "Market efficiency and co-integration: An application to the Sterling and Deutschemark exchange markets," *Journal of International Money and Finance*, 8, 75–88.

[9] Hall, P. (1982). "On some simple estimates of an exponent of regular variation," *Journal of the Royal Statistical Society, Series B*, 44, 37–42.

[10] Hall, P. and A. H. Welsh (1985). "Adaptive estimates of parameters of regular variation," *Annals of Statistics*, 13, 331–341.

[11] Hannan, E. J. and J. Rissanen (1982). "Recursive estimation of ARMA order," *Biometrika*, 69, 273–280 [Corrigenda, *Biometrika*, 1983, 70].

[12] Hill, B. M. (1975). "A simple general approach to inference about the tail of a distribution," *Annals of Statistics*, 3, 1163–1174.

[13] Juttner, J. (1990). *Financial Markets, Interest Rates and Monetary Economics.* Melbourne: Longman Cheshire.

[14] Knight, K. (1989). "Limit theory for autoregressive parameters in an infinite variance random walk," *Canadian Journal of Statistics*, 261–278.

[15] Knight, K. (1991). "Limit theory for M-estimates in an integrated infinite variance process," *Econometric Theory*, 7, 200–212.

[16] Lee, Chin and P. C. B. Phillips (1994). "An improved ARMA-prewhitened estimator of the long-run variance with an application to unit root tests," Yale University, mimeographed.

[17] Loretan, M. and P. C. B. Phillips (1994). "Testing covariance stationarity of heavy tailed time series: An overview of the theory with applications to several financial datasets," *Journal of Empirical Finance*, 1, 211–248.

[18] McFarland, J. W., P. C. McMahon and Y. Ngama (1992). "Forward exchange rates and expectations during the 1920's: A reexamination of the evidence," *Scottish Journal of Political Economy,* 39, 129-140.

[19] Moore, M. J. (1992). "Some new tests of efficiency in the forward exchange market," mimeographed, Queens University of Belfast.

[20] Ouliaris, S. and P. C. B. Phillips (1993). "COINT 2.0: regression software for cointegrated systems," available from Aptech Systems.

[21] Park, J. Y. (1992). "Canonical cointegrating regressions," *Econometrica*, 60, 119–143.

[22] Park, J. Y. and P. C. B. Phillips (1988). "Statistical inference in regressions with integrated processes: Part 1," *Econometric Theory*, 4, 468–497.

[23] Phillips, P. C. B. (1990). "Time series regression with a unit root and infinite variance errors," *Econometric Theory*, 6, 44–62.

[24] Phillips, P. C. B. (1991). "A shortcut to LAD estimator asymptotics," *Econometric Theory*, 7, 450-463.

[25] Phillips, P. C. B. (1995). "Robust nonstationary regression," *Econometric Theory,* 11, 912-951.

[26] Phillips, P. C. B. (1996). "Econometric Model Determination," *Econometra* (forthcoming).

[27] Phillips, P. C. B. and S. N. Durlauf (1986). "Multiple time series regression with integrated variables," *Review of Economic Studies*, 53, 473–496.

[28] Phillips, P. C. B. and B. E. Hansen (1990). "Statistical inference in instrumental variables regression with I(1) processes," *Review of Economic Studies*, 57, 99–125.

[29] Phillips, P. C. B., McFarland, J. W. and P. C. McMahon (1996). "Robust tests of forward exchange market efficiency with empirical evidence from the 1920's," *Journal of Applied Econometrics,* 11, 1 - 23.

[30] Phillips, P. C. B. and J. W. McFarland (1996). "On Uncovered Interest Rate Parity Regressions", (currently in preparation).

[31] Phillips, P. C. B. and S. Ouliaris (1990), "Asymptotic properties of residual based tests for cointegration," *Econometrica*, 58, 165–194.

[32] Phillips, P. C. B. and W. Ploberger (1994). "Posterior odds testing for a unit root with data-based model selection," *Econometric Theory,* 10, 774-808

[33] Phillips, P. C. B. and W. Ploberger (1996). "An Asymptotic Theory of Bayesian Inference for Time Series," *Econometrica*, 64, 381-413.

[34] Schwarz, G. (1978). "Estimating the dimension of a model," *Annals of Statistics*, 6, 461–464.

[35] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.